



PCT09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,918A

DATE: 06/06/2002
TIME: 12:35:17

Input Set : A:\16430seq.txt
Output Set: N:\CRF3\06062002\I673918A.raw

p.6

3 <110> APPLICANT: Xia, Zhi-Qiang
 4 Costa, Michael A
 5 Davin, Laurence B
 6 Lewis, Norman G
 8 <120> TITLE OF INVENTION: Recombinant Secoisolariciresinol Dehydrogenase, and
 9 Methods of Use
 11 <130> FILE REFERENCE: WSUR116430
 13 <140> CURRENT APPLICATION NUMBER: 09/673,918A
 14 <141> CURRENT FILING DATE: 2002-04-09
 16 <150> PRIOR APPLICATION NUMBER: PCT/US99/08975
 17 <151> PRIOR FILING DATE: 1999-04-23
 19 <150> PRIOR APPLICATION NUMBER: 60/082,977
 20 <151> PRIOR FILING DATE: 1998-04-24
 22 <160> NUMBER OF SEQ ID NOS: 25
 24 <170> SOFTWARE: PatentIn Ver. 2.0
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 819
 28 <212> TYPE: DNA
 29 <213> ORGANISM: Forsythia x intermedia
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: (1)..(819)
 35 <400> SEQUENCE: 1
 36 atg cag ctt cga act gca ttc gca aga agg cta gaa gga aaa gtt gcc 48
 37 Met Gln Leu Arg Thr Ala Phe Ala Arg Arg Leu Glu Gly Lys Val Ala
 38 1 5 10 15
 40 ctt ata aca gga gga gcc agt gga att gga gaa acc aca gca aaa ctc 96
 41 Leu Ile Thr Gly Gly Ala Ser Gly Ile Gly Glu Thr Thr Ala Lys Leu
 42 20 25 30
 44 ttc tcc caa cat gga gcc aaa gtt gcc att gct gat gtccaa gat gaa 144
 45 Phe Ser Gln His Gly Ala Lys Val Ala Ile Ala Asp Val Gln Asp Glu
 46 35 40 45
 48 tta ggt cac tca gtt gtc gag gcc att ggc act tcc aat tcc acc tac 192
 49 Leu Gly His Ser Val Val Glu Ala Ile Gly Thr Ser Asn Ser Thr Tyr
 50 50 55 60
 52 atc cac tgt gat gtt act aat gaa gac ggt gtt aaa aat gcc gtg gac 240
 53 Ile His Cys Asp Val Thr Asn Glu Asp Gly Val Lys Asn Ala Val Asp
 54 65 70 75 80
 56 aac aca gtt tca acc tat gga aaa ctg gac att atg ttc agc aat gca 288
 57 Asn Thr Val Ser Thr Tyr Gly Lys Leu Asp Ile Met Phe Ser Asn Ala
 58 85 90 95
 60 gga att tct gat ccc aac agg ccc cgc atc ata gac aac gaa aaa gca 336
 61 Gly Ile Ser Asp Pro Asn Arg Pro Arg Ile Ile Asp Asn Glu Lys Ala

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,918A

DATE: 06/06/2002
TIME: 12:35:17

Input Set : A:\16430seq.txt
Output Set: N:\CRF3\06062002\I673918A.raw

62	100	105	110	
64	gac ttt gaa cgc gtt ctc agt gta aat gta acc gga gtt ttc cta tgc			384
65	Asp Phe Glu Arg Val Leu Ser Val Asn Val Thr Gly Val Phe Leu Cys			
66	115	120	125	
68	atg aag cac gca gca cgt gtt atg att cca gca cgc agt ggc aac ata			432
69	Met Lys His Ala Ala Arg Val Met Ile Pro Ala Arg Ser Gly Asn Ile			
70	130	135	140	
72	att tcc act gct agt tta agc tca act atg ggt ggt ggt tct tca cat			480
73	Ile Ser Thr Ala Ser Leu Ser Ser Thr Met Gly Gly Gly Ser Ser His			
74	145	150	155	160
76	gcc tat tgt ggt tca aag cat gct gtg tta gcc ctt act agg aat ctg			528
77	Ala Tyr Cys Gly Ser Lys His Ala Val Leu Ala Leu Thr Arg Asn Leu			
78	165	170	175	
80	gca gtc gag ctc gga caa ttt ggc att agg gtt aat tgt ttg tct cct			576
81	Ala Val Glu Leu Gly Gln Phe Gly Ile Arg Val Asn Cys Leu Ser Pro			
82	180	185	190	
84	tcc ggg ctt cct acg gct tta ggc aag aaa ttt tca ggg att aaa aat			624
85	Phe Gly Leu Pro Thr Ala Leu Gly Lys Phe Ser Gly Ile Lys Asn			
86	195	200	205	
88	gaa gaa gaa ttt gag aat gta ata aac ttt gcg gga aat ttg aaa ggt			672
89	Glu Glu Phe Glu Asn Val Ile Asn Phe Ala Gly Asn Leu Lys Gly			
90	210	215	220	
92	cca aaa ttt aat gtt gag gat gtt gcc aat gca gct ctt tat ctg gct			720
93	Pro Lys Phe Asn Val Glu Asp Val Ala Asn Ala Ala Leu Tyr Leu Ala			
94	225	230	235	240
96	agt gat gag gca aaa tac gtg agt gga cac aat ctg ttc att gat gga			768
97	Ser Asp Glu Ala Lys Tyr Val Ser Gly His Asn Leu Phe Ile Asp Gly			
98	245	250	255	
100	ggg ttc agc gtc tgc aat tct gta atc aaa gtg ttc caa tat cca gat			816
101	Gly Phe Ser Val Cys Asn Ser Val Ile Lys Val Phe Gln Tyr Pro Asp			
102	260	265	270	
104	tct			819
105	Ser			
108	<210> SEQ ID NO: 2			
109	<211> LENGTH: 273			
110	<212> TYPE: PRT			
111	<213> ORGANISM: Forsythia x intermedia			
113	<400> SEQUENCE: 2			
114	Met Gln Leu Arg Thr Ala Phe Ala Arg Arg Leu Glu Gly Lys Val Ala			
115	1	5	10	15
117	Leu Ile Thr Gly Gly Ala Ser Gly Ile Gly Glu Thr Thr Ala Lys Leu			
118	20	25	30	
120	Phe Ser Gln His Gly Ala Lys Val Ala Ile Ala Asp Val Gln Asp Glu			
121	35	40	45	
123	Leu Gly His Ser Val Val Glu Ala Ile Gly Thr Ser Asn Ser Thr Tyr			
124	50	55	60	
126	Ile His Cys Asp Val Thr Asn Glu Asp Gly Val Lys Asn Ala Val Asp			
127	65	70	75	80
129	Asn Thr Val Ser Thr Tyr Gly Leu Asp Ile Met Phe Ser Asn Ala			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,918A

DATE: 06/06/2002
TIME: 12:35:17

Input Set : A:\16430seq.txt
Output Set: N:\CRF3\06062002\I673918A.raw

130	85	90	95
132	Gly Ile Ser Asp Pro Asn Arg Pro Arg Ile Ile Asp Asn Glu Lys Ala		
133	100	105	110
135	Asp Phe Glu Arg Val Leu Ser Val Asn Val Thr Gly Val Phe Leu Cys		
136	115	120	125
138	Met Lys His Ala Ala Arg Val Met Ile Pro Ala Arg Ser Gly Asn Ile		
139	130	135	140
141	Ile Ser Thr Ala Ser Leu Ser Ser Thr Met Gly Gly Gly Ser Ser His		
142	145	150	155
144	Ala Tyr Cys Gly Ser Lys His Ala Val Leu Ala Leu Thr Arg Asn Leu		
145	165	170	175
147	Ala Val Glu Leu Gly Gln Phe Gly Ile Arg Val Asn Cys Leu Ser Pro		
148	180	185	190
150	Phe Gly Leu Pro Thr Ala Leu Gly Lys Lys Phe Ser Gly Ile Lys Asn		
151	195	200	205
153	Glu Glu Glu Phe Glu Asn Val Ile Asn Phe Ala Gly Asn Leu Lys Gly		
154	210	215	220
156	Pro Lys Phe Asn Val Glu Asp Val Ala Asn Ala Ala Leu Tyr Leu Ala		
157	225	230	235
159	Ser Asp Glu Ala Lys Tyr Val Ser Gly His Asn Leu Phe Ile Asp Gly		
160	245	250	255
162	Gly Phe Ser Val Cys Asn Ser Val Ile Lys Val Phe Gln Tyr Pro Asp		
163	260	265	270
165	Ser		
169	<210> SEQ ID NO: 3		
170	<211> LENGTH: 831		
171	<212> TYPE: DNA		
172	<213> ORGANISM: Forsythia x intermedia		
174	<220> FEATURE:		
175	<221> NAME/KEY: CDS		
176	<222> LOCATION: (1)..(831)		
178	<400> SEQUENCE: 3		
179	atg gca gcc act tca cag gtt cta act gca atc gca aga agg cta gaa	48	
180	Met Ala Ala Thr Ser Gln Val Leu Thr Ala Ile Ala Arg Arg Leu Glu		
181	1 5 10 15		
183	gga aaa gtt gcc ctt ata aca gga gga gcc agt gga att gga gaa acc	96	
184	Gly Lys Val Ala Leu Ile Thr Gly Gly Ala Ser Gly Ile Gly Glu Thr		
185	20 25 30		
187	aca gca aaa ctc ttc tcc caa cat gga gcc aaa gtt gcc att gct gat	144	
188	Thr Ala Lys Leu Phe Ser Gln His Gly Ala Lys Val Ala Ile Ala Asp		
189	35 40 45		
191	gtc caa gat gaa tta ggt cac tca gtt gtc gag gcc att ggc act tcc	192	
192	Val Gln Asp Glu Leu Gly His Ser Val Val Glu Ala Ile Gly Thr Ser		
193	50 55 60		
195	aat tcc acc tac atc cac tgt gat gtt act aat gaa gac ggt gtt aaa	240	
196	Asn Ser Thr Tyr Ile His Cys Asp Val Thr Asn Glu Asp Gly Val Lys		
197	65 70 75 80		
199	aat gcc gtg gac aac aca gtt tca acc tat gga aaa ctg gac att atg	288	
200	Asn Ala Val Asp Asn Thr Val Ser Thr Tyr Gly Lys Leu Asp Ile Met		

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,918A

DATE: 06/06/2002
TIME: 12:35:17

Input Set : A:\16430seq.txt
Output Set: N:\CRF3\06062002\I673918A.raw

201	85	90	95	
203	ttc agc aat gca gga att tct gat ccc aac agg ccc cgc atc ata gac			336
204	Phe Ser Asn Ala Gly Ile Ser Asp Pro Asn Arg Pro Arg Ile Ile Asp			
205	100	105	110	
207	aac gaa aaa gca gac ttt gaa cgc gtt ttc agt gta aat gta acc gga			384
208	Asn Glu Lys Ala Asp Phe Glu Arg Val Phe Ser Val Asn Val Thr Gly			
209	115	120	125	
211	gtt ttc cta tgc atg aag cac gca gca cgt gtt atg att cca gca cgc			432
212	Val Phe Leu Cys Met Lys His Ala Ala Arg Val Met Ile Pro Ala Arg			
213	130	135	140	
215	agt ggc aac ata att tcc act gct agt tta agc tca act atg ggt ggt			480
216	Ser Gly Asn Ile Ile Ser Thr Ala Ser Leu Ser Ser Thr Met Gly Gly			
217	145	150	155	160
219	ggt tct tca cat gcc tat tgt ggt tca aag cat gct gtg tta ggc ctt			528
220	Gly Ser Ser His Ala Tyr Cys Gly Ser Lys His Ala Val Leu Gly Leu			
221	165	170	175	
223	act agg aat ctg gca gtc gag ctc gga caa ttt ggc att agg gtt aat			576
224	Thr Arg Asn Leu Ala Val Glu Leu Gly Gln Phe Gly Ile Arg Val Asn			
225	180	185	190	
227	tgt ttg tct ctc ggg ctt cct acg gct tta ggc aag aaa ttt tca			624
228	Cys Leu Ser Pro Phe Gly Leu Pro Thr Ala Leu Gly Lys Lys Phe Ser			
229	195	200	205	
231	ggg att aaa aat gaa gaa gaa ttt gag aat gta ata aac ttt gcg gga			672
232	Gly Ile Lys Asn Glu Glu Glu Phe Glu Asn Val Ile Asn Phe Ala Gly			
233	210	215	220	
235	aat ctg aaa ggt cca aaa ttt aat gtt gag gat gtt gcc aat gca gct			720
236	Asn Leu Lys Gly Pro Lys Phe Asn Val Glu Asp Val Ala Asn Ala Ala			
237	225	230	235	240
239	ctt tat ctg gct agt gat gag gca aaa tac gtg agt gga cac aat ctg			768
240	Leu Tyr Leu Ala Ser Asp Glu Ala Lys Tyr Val Ser Gly His Asn Leu			
241	245	250	255	
243	ttc att gat gga ggg ttc agc gtc tgc aat tct gta atc aaa gtg ttc			816
244	Phe Ile Asp Gly Gly Phe Ser Val Cys Asn Ser Val Ile Lys Val Phe			
245	260	265	270	
247	caa tat cca gat tct			831
248	Gln Tyr Pro Asp Ser			
249	275			
252	<210> SEQ ID NO: 4			
253	<211> LENGTH: 277			
254	<212> TYPE: PRT			
255	<213> ORGANISM: Forsythia x intermedia			
257	<400> SEQUENCE: 4			
258	Met Ala Ala Thr Ser Gln Val Leu Thr Ala Ile Ala Arg Arg Leu Glu			
259	1	5	10	15
261	Gly Lys Val Ala Leu Ile Thr Gly Gly Ala Ser Gly Ile Gly Glu Thr			
262	20	25	30	
264	Thr Ala Lys Leu Phe Ser Gln His Gly Ala Lys Val Ala Ile Ala Asp			
265	35	40	45	
267	Val Gln Asp Glu Leu Gly His Ser Val Val Glu Ala Ile Gly Thr Ser			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,918A

DATE: 06/06/2002
TIME: 12:35:17

Input Set : A:\16430seq.txt
Output Set: N:\CRF3\06062002\I673918A.raw

268	50	55	60	
270	Asn Ser Thr Tyr Ile His Cys Asp Val Thr Asn Glu Asp Gly Val Lys			
271	65	70	75	80
273	Asn Ala Val Asp Asn Thr Val Ser Thr Tyr Gly Lys Leu Asp Ile Met			
274	85	90	95	
276	Phe Ser Asn Ala Gly Ile Ser Asp Pro Asn Arg Pro Arg Ile Ile Asp			
277	100	105	110	
279	Asn Glu Lys Ala Asp Phe Glu Arg Val Phe Ser Val Asn Val Thr Gly			
280	115	120	125	
282	Val Phe Leu Cys Met Lys His Ala Ala Arg Val Met Ile Pro Ala Arg			
283	130	135	140	
285	Ser Gly Asn Ile Ile Ser Thr Ala Ser Leu Ser Ser Thr Met Gly Gly			
286	145	150	155	160
288	Gly Ser Ser His Ala Tyr Cys Gly Ser Lys His Ala Val Leu Gly Leu			
289	165	170	175	
291	Thr Arg Asn Leu Ala Val Glu Leu Gly Gln Phe Gly Ile Arg Val Asn			
292	180	185	190	
294	Cys Leu Ser Pro Phe Gly Leu Pro Thr Ala Leu Gly Lys Lys Phe Ser			
295	195	200	205	
297	Gly Ile Lys Asn Glu Glu Glu Phe Glu Asn Val Ile Asn Phe Ala Gly			
298	210	215	220	
300	Asn Leu Lys Gly Pro Lys Phe Asn Val Glu Asp Val Ala Asn Ala Ala			
301	225	230	235	240
303	Leu Tyr Leu Ala Ser Asp Glu Ala Lys Tyr Val Ser Gly His Asn Leu			
304	245	250	255	
306	Phe Ile Asp Gly Gly Phe Ser Val Cys Asn Ser Val Ile Lys Val Phe			
307	260	265	270	
309	Gln Tyr Pro Asp Ser			
310	275			
313	<210> SEQ_ID NO: 5			
314	<211> LENGTH: 819			
315	<212> TYPE: DNA			
316	<213> ORGANISM: Forsythia x intermedia			
318	<220> FEATURE:			
319	<221> NAME/KEY: CDS			
320	<222> LOCATION: (1)..(819)			
322	<220> FEATURE:			
323	<221> NAME/KEY: misc_feature			
324	<222> LOCATION: (1)..(819)			
325	<223> OTHER INFORMATION: cDNA molecule encoding secoisolariciresinol dehydrogenase			
wherein Xaa = any amino acid				
327	<400> SEQUENCE: 5			
328	atg cag ctt cga act gca atc gca aga agg cta gaa gga aaa gtt gcc	48		
329	Met Gln Leu Arg Thr Ala Ile Ala Arg Arg Leu Glu Gly Lys Val Ala			
330	1	5	10	15
332	ctt ata aca gga gga gcc agt gga gtt gga gaa gtc aca gca aaa ctc	96		
333	Leu Ile Thr Gly Gly Ala Ser Gly Val Gly Glu Val Thr Ala Lys Leu			
334	20	25	30	
336	tcc tcc caa cat gga gcc aaa gtt gcc att gct gat gtc caa gat gaa	144		
337	Phe Ser Gln His Gly Ala Lys Val Ala Ile Ala Asp Val Gln Asp Glu			

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 06/06/2002
 PATENT APPLICATION: US/09/673,918A TIME: 12:35:18

Input Set : A:\16430seq.txt
 Output Set: N:\CRF3\06062002\I673918A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 360,366
 Seq#:5; Xaa Pos. 120,122
 Seq#:6; Xaa Pos. 120,122
 Seq#:11; Xaa Pos. 3,12,20
 Seq#:14; N Pos. 3,9,15,18
 Seq#:15; N Pos. 3,9
 Seq#:16; N Pos. 3,9
 Seq#:18; N Pos. 3,9

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23
 Seq#:1; Line(s) 24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43
 Seq#:1; Line(s) 44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63
 Seq#:1; Line(s) 64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83
 Seq#:1; Line(s) 84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102
 Seq#:1; Line(s) 103,104,105,106,107,108
 Seq#:2; Line(s) 109,110,111,112,113,114,115,116,117,118,119,120,121,122,123
 Seq#:2; Line(s) 124,125,126,127,128,129,130,131,132,133,134,135,136,137,138
 Seq#:2; Line(s) 139,140,141,142,143,144,145,146,147,148,149,150,151,152,153
 Seq#:2; Line(s) 154,155,156,157,158,159,160,161,162,163,164,165,166,167,168
 Seq#:2; Line(s) 169
 Seq#:3; Line(s) 170,171,172,173,174,175,176,177,178,179,180,181,182,183,184
 Seq#:3; Line(s) 185,186,187,188,189,190,191,192,193,194,195,196,197,198,199
 Seq#:3; Line(s) 200,201,202,203,204,205,206,207,208,209,210,211,212,213,214
 Seq#:3; Line(s) 215,216,217,218,219,220,221,222,223,224,225,226,227,228,229
 Seq#:3; Line(s) 230,231,232,233,234,235,236,237,238,239,240,241,242,243,244
 Seq#:3; Line(s) 245,246,247,248,249,250,251,252
 Seq#:4; Line(s) 253,254,255,256,257,258,259,260,261,262,263,264,265,266,267
 Seq#:4; Line(s) 268,269,270,271,272,273,274,275,276,277,278,279,280,281,282
 Seq#:4; Line(s) 283,284,285,286,287,288,289,290,291,292,293,294,295,296,297
 Seq#:4; Line(s) 298,299,300,301,302,303,304,305,306,307,308,309,310,311,312
 Seq#:4; Line(s) 313
 Seq#:5; Line(s) 314,315,316,317,318,319,320,321,322,323,324,325,326,327,328
 Seq#:5; Line(s) 329,330,331,332,333,334,335,336,337,338,339,340,341,342,343
 Seq#:5; Line(s) 344,345,346,347,348,349,350,351,352,353,354,355,356,357,358
 Seq#:5; Line(s) 359,360,361,362,363,364,365,366,367,368,369,370,371,372,373
 Seq#:5; Line(s) 374,375,376,377,378,379,380,381,382,383,384,385,386,387,388
 Seq#:5; Line(s) 389,390,391,392,393,394,395,396,397,398,399,400
 Seq#:6; Line(s) 401,402,403,404,405,406,407,408,409,410,411,412,413,414,415
 Seq#:6; Line(s) 416,417,418,419,420,421,422,423,424,425,426,427,428,429,430
 Seq#:6; Line(s) 431,432,433,434,435,436,437,438,439,440,441,442,443,444,445